

SPEECH SYNTHESIS SYSTEM

2

Patent Number: JP9185393
Publication date: 1997-07-15
Inventor(s): KOGURE YOSHIYA
Applicant(s): NEC CORP
Requested Patent: JP9185393
Application Number: JP19950353395 19951228
Priority Number(s):
IPC Classification: G10L5/04; G10L3/00
EC Classification:
Equivalents:

Abstract

PROBLEM TO BE SOLVED: To synthesize an arbitrary word which is almost natural by registering speech data in 1-syllable units and speech data spoken in daily-use KANJI (Chinese character) units and properly selecting and combining them together.

SOLUTION: This system selects (n) syllables out of speech data generated by plural Japanese syllables, syllable by syllable, and plural syllables and registers speech data, spoken in (n)-syllable units (n: integer larger than 2) by all combinations, on a storage medium. For speech synthesis, the speech data are combined and plural speech data spoken with several kinds of intonation (speech pitch) for one kind of syllable are registered as speech data in 1-syllable units. For speech data in (n)-syllable units, plural speech data spoken while the intonation position is changed are registered for respective (n) syllables.

Data supplied from the **esp@cenet** database - I2

(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平9-185393

(43)公開日 平成9年(1997)7月15日

(51)Int.Cl.⁶G 1 0 L 5/04
3/00

識別記号

府内整理番号

F I

G 1 0 L 5/04
3/00

技術表示箇所

F
H

審査請求 有 請求項の数5 FD (全5頁)

(21)出願番号 特願平7-353395

(22)出願日 平成7年(1995)12月28日

(71)出願人 000004237

日本電気株式会社

東京都港区芝五丁目7番1号

(72)発明者 小暮 佳也

東京都港区芝五丁目7番1号 日本電気株式会社内

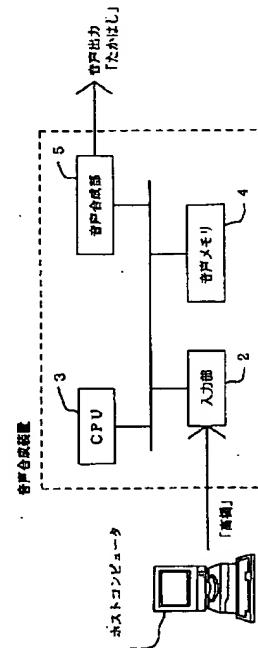
(74)代理人 弁理士 加藤 朝道

(54)【発明の名称】 音声合成方式

(57)【要約】

【課題】1音節単位の音声データと常用漢字単位で発声した音声データを登録し、適当に選択し組み合わせ合成することにより自然に近い任意の単語を合成する。

【解決手段】日本語の複数ある音節を各々1音節単位で発声した音声データと複数の音節の中からn音節を選び、全ての組み合わせでn音節単位(nは2以上の整数)で発声した音声データを記憶媒体に登録し、音声合成の際に、音声データを組み合わせ、任意単語の音声波形を合成するものであり、1音節単位の音声データは、1種類の音節に対し、数種類のアクセント(アクセントは音声ピッチの高低)で発音した複数の音声データを登録する。n音節単位の音声データは、各n音節に対し、アクセントを付ける位置を変えて発声した複数の音声データを登録する。



【特許請求の範囲】

【請求項1】録音再生型の音声合成装置において、日本語の複数の音節を各々1音節単位で発声してなる音声データと、前記複数の音節のうちのn音節について所定の組み合わせにてn音節単位（但し、nは2以上の所定の整数）で発声してなる音声データと、を記憶手段に予め登録し、音声合成の際に、前記記憶手段に登録された前記音声データを組み合わせて任意の単語の音声波形を合成するようにしたことを特徴とする音声合成方式。

【請求項2】前記1音節単位の音声データについては、1種類の音節に対し複数種のアクセント（但し、アクセントは音声ピッチの高低）で発声してなる複数の音声データを登録し、

前記n音節単位の音声データについては、各n音節に対し、アクセントを付ける位置を変えて発声してなる複数の音声データを登録することを特徴とする請求項1記載の音声合成方式。

【請求項3】前記1音節及び／又は前記n音節単位の音声データが、常用漢字単位で用いられる音節に基づいて作成されていることを特徴とする請求項1又は2記載の音声合成方式。

【請求項4】前記1音節単位の音声データについてはアクセントの有無に応じて互いに異なる識別コードが付され、前記n音節単位の音声データについては、全音節アクセント無し及びアクセントが付される音節の位置に応じて、互いに異なる識別コードが付され、前記音声データが前記識別コードと対応付けられて前記記憶手段に登録されることを特徴とする請求項2記載の音声合成方式。

【請求項5】任意の単語を音声合成する際に、前記任意の単語を1音節単位及び／又はn音節単位の列に分解し、前記1音節については前記単語の発声時の前記1音節のアクセントの有無に基づき前記識別コードを選択し、前記n音節については前記単語の発声時の前記n音節のアクセント無し又はアクセントの位置に応じて前記識別コードを選択し、前記選択された識別コードを編集してなる識別コード列情報をに基づき前記記憶手段から対応する音声データを読み出し前記任意の単語の音声を合成することを特徴とする請求項4記載の音声合成方式。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、日本語の音声を合成する方式に関し、特に録音再生型音声合成装置において、登録した音声を組み合わせ任意の単語を合成する音声合成方式に関する。

【0002】

【従来の技術】従来の録音再生型音声合成装置においては、通常、文章を、文節または単語毎に区切って記憶媒

体に予め登録しておき、再生（音声合成）時に、これらの文節又は単語を組み合わせることにより音声波形を合成している。

【0003】そして、特に、人名や地名等のように、全てを登録することが困難な単語については、1音節単位（あ、い、う、…）の音声データを予め登録しておき、1音節単位の音声データを組み合わせて列ねることにより、単語を合成している。

【0004】また、日本語のアクセントはピッチの変化であることから、組み合わせた単語にアクセントを付けるために、1音節につき複数のピッチ（低、高等）の音節データを予め登録しておき、合成する単語に応じて、適当なピッチの音節データを選択し、これらの1音節単位の音声データを組み合わせて音声波形を合成している。

【0005】

【発明が解決しようとする課題】しかしながら、従来の録音再生型音声合成装置において、人名や地名等を1音節単位の音声データを組み合わせて音声合成した場合、たとえ、複数のピッチのデータの中から適切なピッチの音節データを選択して組み合わせても、各音節がそれぞれ独立に発声された単位であるために、連続性の無い不自然な音声波形が合成出力されてしまうという問題点を有している。

【0006】従って、本発明は、上記従来技術の問題点に鑑みて為されたものであって、1音節単位の音声データと好ましくは常用漢字単位で発声した音声データを登録しておき、これらを適宜選択し組み合わせて合成することにより、自然に近い任意の単語を合成することを可能とする音声合成方式を提供することを目的とする。

【0007】

【課題を解決するための手段】前記目的を達成するため、本発明は、録音再生型の音声合成装置において、日本語の複数の音節を各々1音節単位で発声してなる音声データと、前記複数の音節のうちのn音節について所定の組み合わせにてn音節単位（但し、nは2以上の所定の整数）で発声してなる音声データと、を記憶手段に予め登録し、音声合成の際に、前記記憶手段に登録された前記音声データを組み合わせて任意の単語の音声波形を合成するようにしたことを特徴とする音声合成方式を提供する。

【0008】また、本発明は、前記1音節単位の音声データについては、1種類の音節に対し複数種のアクセント（但し、アクセントは音声ピッチの高低）で発声してなる複数の音声データを登録し、前記n音節単位の音声データについては、各n音節に対し、アクセントを付ける位置を変えて発声してなる複数の音声データを登録することを特徴としたものである。

【0009】そして、本発明においては、好ましくは、前記第1音節及び前記n音節単位の音声データは常用漢

字単位で用いられる音節に基づいて作成されている。具体的には、例えば、地名や人名等の漢字1文字に対応する発声を1データとして登録する。

【0010】

【発明の実施の形態】本発明の実施の形態について図面を参照して以下に詳細に説明する。

【0011】図1は、本発明の一実施形態の全体の構成を示す図であり、本発明に係る音声合成装置の典型的な構成を示した図である。図2は、本発明の一実施形態に係る音声合成方式における音声メモリのデータ構造の一例を示した図である。図3は、本発明の一実施形態による音声合成の一例を模式的に示した図である。

【0012】図1を参照して、音声合成装置は、ホストコンピュータ1からデータを入力する入力部2、CPU3、音声メモリ4、及び音声合成部5からなる。

【0013】音声メモリ4には、1音節単位(あ、い、う、….)の音声データと、2音節単位(あい、いう、うえ、….)の音声データと、が記憶格納されている。

【0014】そして、音声データ毎に、互いに異なる識別番号(1D番号；音声データを一義的に識別するための番号)が付されている(図2参照)。

【0015】1音節単位の音声データは、ある一つの音節に対し、アクセントのある音声データと、アクセントの無い音声データとがある。1組の2音節(例えば、「あい」)に対しては、前の音節にアクセントのある組み合わせ、後ろの音節にアクセントのある組み合わせ、そしてアクセントの無い組み合わせの3種類の音声データを格納する。

【0016】すなわち、図2を参照して、一音節「み」について、アクセントの有無に応じて識別番号は“0123”、“0124”とされ、また2音節「たか」について、前の音節「た」にアクセントがあるものは識別番号“0500”、後ろの音節「か」にアクセントがあるものとして識別番号“0501”、アクセントがない組み合わせについては、識別番号“0502”が付与され、対応する音声データと関連させて音声メモリ4に格納されている。本実施形態においては、1音節、2音節単位の音声データは、好ましくは、常用漢字単位で用いられる音節に基づいて作成される。

【0017】ある単語の音声を合成する際には、まずホストコンピュータ1で予め編集された識別番号の列の情報が入力部2に入力される。

【0018】CPU3は入力された識別番号に対応する音声データ(ディジタル信号)を音声メモリ4から読み

出して音声合成部5に供給し、音声合成部5はアナログ音声波形に変換出力して合成音声を出力する。

【0019】次に、音声データを組み合わせる方法について説明する。

【0020】図3に示すように、人名の「高橋(たかはし)」を合成する場合、「か」にアクセントのある2音節「たか」と、アクセントの無い「はし」を組み合わせると、自然である。

【0021】同じく、「三鷹(みたか)」を合成する場合は、アクセントの無い1音節「み」と、アクセントの無い2音節「たか」を組み合わせればよい。

【0022】日本語の1音節の種類は「あ、い、う、…」に加え、「が、だ、ば、…」の濁音、「ぱ、ぴ、…」の半濁音、「きや、びい、ぴや、…」の拗音などを含め合計で百数十種類であり、上記実施形態のように、全て2音節の組み合わせは、これらの組み合わせで約1万種類である。

【0023】このため、2音節のある1つの組み合わせについて3種類のアクセントの音声データとすると、合計約3万種類の音声を音声メモリに格納することによって、録音再生型の音声合成装置でも、任意の音声をより自然に近い発音で音声波形を合成することが可能となる。

【0024】

【発明の効果】以上説明したように、本発明によれば、録音再生型音声合成装置に1音節単位の音声データと、前記1音節単位の音声データを組み合わせたn音節単位の音声データとを登録し、適当なアクセントの音声データを組み合わせて呼び出し、任意単語の音声波形を合成することによって、任意の単語をより自然に近い音声で合成することができる。

【図面の簡単な説明】

【図1】本発明の一実施形態に係る録音再生型音声合成装置の構成を示す図である。

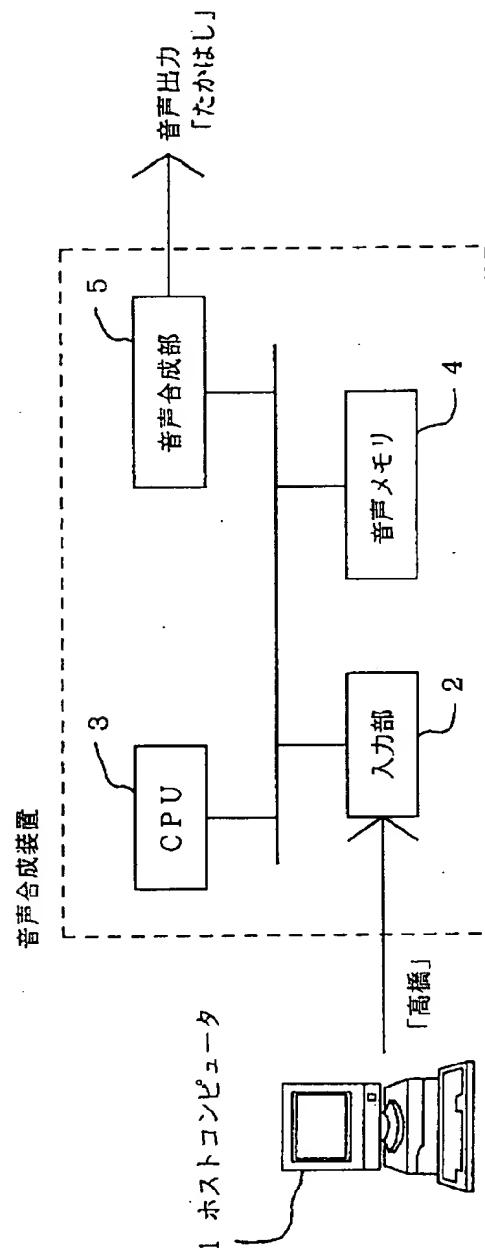
【図2】本発明の一実施形態における音声メモリのデータ格納の状態を示す図である。

【図3】本発明の一実施形態による単語の音声合成の様子を模式的に示す図である。

【符号の説明】

- 1 ホストコンピュータ
- 2 入力部
- 3 CPU
- 4 音声メモリ
- 5 音声合成部

【図1】



【図2】

識別番号	音声内容
0 1 2 3	み
0 1 2 4	み
0 5 0 0	た か
0 5 0 1	た か
0 5 0 2	た か

注：^一はアクセント位置を示す。

【図3】

高 橋 →	た	か	・	は	し
三 魔 →	み	・	た	か	